NAS12-554

*AD*/*COM*

INTERIM SCIENTIFIC REPORT

DEVELOPMENT OF ADVANCED DIGITAL TECHNIQUES FOR
DATA ACQUISITION PROCESSING AND COMMUNICATION

By J. Levy, E.H. Gavenman and W.R. Bechtold

1 November 1967

Prepared under Contract No. NAS12-554(DO-C9)

National Aeronautics and Space Administration
Electronics Research Center
KC/Computer Research Laboratory
Cambridge, Massachusetts

## ADVANCED COMMUNICATIONS

## INFORMATION MANAGEMENT

**RESEARCH**

**DEVELOPMENT**

**ENGINEERING**

## ADCOM, INC.

Benjamin Goldstein and Robert M. Snow
Co-technical Monitors
NAS12-554
575 Technology Square
Cambridge, Massachusetts 02139

# INTERIM SCIENTIFIC REPORT FOR THE
# DEVELOPMENT OF ADVANCED DIGITAL TECHNIQUES FOR
# DATA ACQUISITION PROCESSING AND COMMUNICATION

By J. Levy, E.H. Gavenman and W.R. Bechtold

1 November 1967

# TABLE OF CONTENTS

# TABLE OF CONTENTS (Cont)

## LIST OF ILLUSTRATIONS

v

LIST OF TABLES

INTERIM SCIENTIFIC REPORT
FOR THE DEVELOPMENT OF ADVANCED DIGITAL TECHNIQUES
FOR DATA ACQUISITION PROCESSING AND COMMUNICATION

By J. Levy, E. H. Gavenman and W. R. Bechtold

ADCOM, Inc.
Cambridge, Massachusetts

SUMMARY


This report reviews progress in the following four study areas:

(1) Characterization of Image Data Sources. Descriptions of study methodology, classification of imaging missions, imaging sensors, and image fidelity criteria are presented.

(2) Second Order Interpolation. A discussion of second order interpolation, and a presentation of the first pictorial material compressed thereby are contained in this report.

(3) Sensitivity of Compressed Pictures to Errors. The first simulation subjecting compressed picture data to bit errors have been completed. The effect of bit errors upon a picture compressed in accordance with zero order interpolation is displayed; the effect of the same noise upon the picture when a simple line-substitution error control scheme is employed is also displayed.

(4) Simulation Test of A/D Mapping Technique. A novel A/D mapping technique assigns $\ell$-bit sequences to each of the $2^\ell$ quantized signal levels so as to cause single bit errors to translate into large quantum level displacements. The theory that this technique causes channel errors to manifest themselves as "wild" data points which may be corrected in a run of correlated data has been tested by a computer simulation. The reliability with which wild points may be smoothed close to correct values is seen to be sensitive to the degree of sample-to-sample correlation.

# INTRODUCTION

## General

This report covers work performed under Contract NAS12-554(DO-C9), Development of Advanced Digital Techniques for Data Acquisition Processing and Communication. The objectives pursued under this contract are:

The characterization of image data sources, the development of new compression algorithms (second order and exponential interpolators), and evaluation of the sensitivity of compressed data to various error rates, and an evaluation of various error control techniques for reducing data sensitivity to errors.

Summary of work during reporting period. --(a) A description of the progress made in the task of characterization of image data sources appears in the section entitled Source Characterization. The approach and methodology are laid out in sections entitled General and Study Methodology. The various possible classes and characteristics of imaging missions are treated in the section entitled Mission Definition, while descriptions of the characteristics of image sensors may be found in the section entitled Imaging Sensors. Image fidelity criteria are dealt within the section entitled Fidelity Criteria.

(b) The section entitled Second Order Interpolation Algorithm contains a discussion of second order interpolation, and presents the first results of applying this data compression technique to pictorial material.

(c) The results of the first simulations subjecting compressed picture data to bit errors of the type generated by a binary symmetric channel are to be found in the section entitled Simulation of Effects of Channel Errors Upon Compressed Pictures. The picture chosen was compressed in accordance with the zero order interpolation procedure; simulation of the effects of the same channel noise disturbance upon the same picture when a simple line substitution technique is used for error control is shown to substantially improve picture quality.

(d) A previously described mapping technique assigns $\ell$-bit sequences to each of the $2^\ell$ quantized signal intensity levels so as to cause single bit errors to translate into large intensity level displacements. The section entitled Simulation Test of A/D Mapping Technique presents a simulation run involving pseudorandom correlated data to test the theory that this technique causes channel errors to manifest themselves as "wild" data points between samples which may, through interpolation, be restored to correct or near-correct value. The technique is found to have some merit, but it is quite sensitive to the characteristics of sample-to-sample correlations.

## ANALYSIS

### Source Characterization

General. --The objective of source characterization is to represent the data source by the smallest number of parameters required for determining the most efficient methods of data processing and transmission. The relevant source parameters are those appearing in the expressions for various system design and performance factors. In general, a data source may be classified as continuous or discrete. The output of a continuous source is a continuously variable waveform whose instantaneous fluctuations convey the information. The output of a discrete source consists of a series of selections from a finite number of states. Imaging data sources are generally continuous sources; thus the following discussions are restricted to considerations of continuous source characterization. References 1 through 8 provide a bibliography of technical information on image sources and their characterization.

Information sources can be represented mathematically be random processes; i. e., the output waveform as a function of time can be described in probabilistic terms. For example, the instantaneous value of the source output at any time may be described by the probability with which it can be expected to fall within a specified range of values. If the value of the source output can be predicted with certainty for any arbitrary instant of time, the source is completely deterministic and it delivers no information. The information gained from one instant to the next is determined by the uncertainty about the source output at the later instant, given the value of the source output at the earlier instant.

The information rate of a source is a parameter of fundamental interest in data handling system design. However, it should be noted that the rate of information transfer for a continuous data source depends not only on the source characteristics, but also on the user capability or requirement to resolve source states expressed in terms of a fidelity criterion. The entire space of the continuum of instantaneous values that can be assumed by the source output can be subdivided (quantized) into cells of nonzero dimensions on the basis of the user resolution capability or requirement. Conditional probabilities can be defined for the occupancy of each cell at any instant of time, conditioned on the sample values at previous observation times. The information rate is determined by the average information derivable at each observation instant and by the number of observations per second. The important concept to note is that "source" characterization is somewhat of a misnomer; "source/user pair" characterization may be more meaningful.

Another pertinent point which must be mentioned is that the "imaging data source" output characteristics are determined not only by the transducer parameters but also by the imaged "scene." For example, a transducer may exhibit a dynamic range capability of 100:1, but the output signal may utilize only 20% of this capability due to scene illumination, lack of distinguishable topographical features, or other factors. This phenomenon occurred in the Mariner IV photos of Mars.

Study methodology. --Figure 1 illustrates the basic source characterization study approach. The study commences with a classification of potential missions from which one can define sets of possible imaging experiments in terms of spectral bands and likely ground resolution requirements for specific types of sensors (e.g., radiometers, television, photographic, etc.). The next phase of the study is concerned with classification of sensors which might be used for each experiment (e.g., dielectric (D/E) tape camera, bolometers, vidicons, etc.).
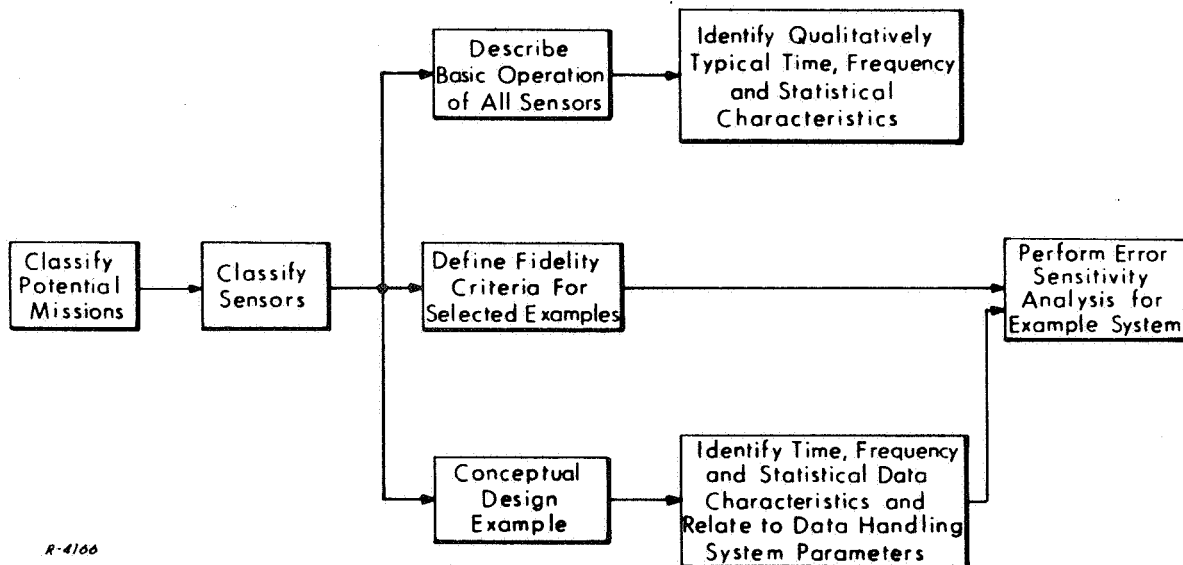
Figure 1  Source Characterization Study Approach

For each mission/sensor combination a conceptual design of the experiment system can be generated including characteristics of the radiation collector (antenna or optics), sensor, and recorder. The system design equations will define the important design parameters of the system including such external parameters as atmospheric turbulence and illumination levels. Due to the magnitude of this task only one experiment system will be analyzed in detail to illustrate the source characterization methodology. Other experiment systems will be defined only in sufficient depth to provide a basic understanding of system operation.

For the example system the time, frequency and probabilistic domain parameters which characterize the output data waveform must be identified. These source parameters then can be related to the data handling system design parameters (e.g., sample rate, quantization, filter characteristics, etc.).

The next step is to define a fidelity criterion for the example system. System performance has in the past been evaluated in terms of the user's subjective judgment of the received picture; there is no general analytic procedure known for relating data fidelity criteria to the fidelity of rendering of the specific kinds of picture detail most essential in a particular picture-taking mission.

It is certainly far beyond the scope of this study to attempt to investigate in great detail the possible fidelity criteria and their analytic relationships to image characteristics. One must instead select a meaningful (if not necessarily optimum) fidelity measure and analyze it.

A fidelity measure, $\Phi$, is a scalar quantity related to the actual data $x(t)$ and the reproduced data $\hat{x}(t)$ over some specified observation interval. It is usually a function of the error between the true data and the data obtained.

$$\Phi \equiv F\left[\hat{x}(t) - x(t)\right] = F\left[e(t)\right] \tag{1}$$

where the functional operation $F$ upon the error between actual and reproduced data describes the fidelity criterion.

The choice of a meaningful fidelity criterion depends strongly on the user's interpretation methods and techniques as well as the experiment design. Thus, the definition of the widest latitude, but meaningful fidelity criterion, is a very difficult task.

The final task of the study is an error sensitivity analysis, i.e., descriptions of error sources in terms of the selected fidelity criterion. These errors may be either systematic or random in nature. Examples of the former are gain changes, certain nonlinearities and zero offset. System noise is an obvious example of the random type of error source. The purpose of this task is to relate the error source to the system design parameters in order to indicate parameter interrelationships which can affect error allocations.

Mission definition. --Four basic planetary mission classes have been identified in this study:

(1)  large area mapping missions,
(2)  small area mapping missions,
(3)  surface (or lander) imaging missions, and
(4)  meteorological missions.

The purpose of the first mission class would be the determination of the gross physical characteristics of the planetary surface including topography, color, and local thermal conditions. Missions of this class could utilize either fly-by or orbiting vehicles of the Mariner or Voyager type, respectively. Spectral bands employed by the imaging experiment would certainly include portions of the microwave and infrared (IR) spectrum. However, the use of the visible spectrum is highly dependent upon the expected cloud cover characteristics of a given planet; visual mapping of planets other than Mars may be impossible.

A mission of the second class would arise upon observation of the results of a large area mapping mission; i.e., specific areas would be selected for high resolution coverage. Typical goals of the small area mapping missions would include determination of soil properties, volcanic activity, faults, and lineaments. These missions could employ orbital vehicles typified by Lunar Orbiter or entry vehicles intended to survive impact. Spectral bands utilized on the orbital vehicles would include microwave, IR and visible. Entry vehicles probably would not use the microwave spectrum.

The third class of missions would be restricted to soft-landing vehicles. The objective of imaging experiments would include site survey (i.e., local topography, color, thermal conditions and soil structure); petrography (rock classification, rock texture, bulk mineralogy); mineralogy (soil composition, hardness, etc.); and life detection (microorganisms). The visible spectrum could be employed for any of the above mission objectives. The infrared spectrum might be useful for site survey. The ultraviolet (UV)

spectral band might be used to detect the presence of certain fluorescent minerals and organic substances. Petrography and mineralogy using the X-ray portion of the spectrum will probably be performed on some missions. It is unclear whether these experiments are to be classified as imaging experiments.

The fourth mission class includes meteorological experiments on an orbital vehicle. Data obtained by such missions would include atmospheric composition and temperature characteristics, cloud movement, cloud structure, turbulence and air currents. Microwave, IR and visible spectral bands will be used to obtain the desired data. In addition to the orbital vehicle experiments, most fly-by vehicles also will carry meteorological experiments. However, these experiments will not produce data on atmospheric motion; main emphasis will be on atmospheric composition, cloud structure and gross temperature characteristics.

The image resolution will differ for each mission class and spectral band within a mission class. In general, these differences may be attributed to physical limitations such as vehicle altitude, antenna size, experiment weight, recorder capacity and smearing due to vehicle motion rather than experimenter requirements. Table I presents a summary of the preceding discussion on imaging mission classification. Included in the table is a column presenting the estimated resolution goals for the various missions and spectral bands. These estimates are based primarily on exploratory missions to Mars or Venus. Resolution estimates for missions to the larger planets (e.g., Jupiter) would probably be several hundred times poorer.

The "Remarks" column of Table I provides an indication of the types of imaging systems which might be used on the various missions. The five systems listed can be summarized as follows:

    (1)   Radiometers. --Single element passive detectors which produce images by mechanical scanning techniques.

    (2)   Side-looking radar. --An active system which produces images via a two-dimensional, multiple pulse correlation technique. This processing is equivalent to generation of a synthetic antenna aperture several miles in length.

    (3)   Television. --Passive detector "array" producing images by electronic scanning techniques.

    (4)   Photography. --Images produced and stored by exposure of film to illumination from the image scene.

## TABLE I

## IMAGING MISSION CLASSIFICATION

| Mission | Description | Spectral Bands | Est. Resolution Goals | Remarks |
|---|---|---|---|---|
| Large Area Mapping | Gross Physical Characteristics | µwave | 10 km | Radiometer |
| | topography | | 100 feet | Side-looking Radar |
| | | IR | 10 km | Radiometer |
| | color | | 100-200 feet | TV; No photos due to film volume requirements |
| | temperature | Visible | 50-100 feet | TV; D/E Tape Camera; No photos |
| Small Area Mapping | High Resolution exploration of limited areas | µwave | 20-50 feet | Side-looking Radar (Orbital Mission Only) |
| | | IR | 10-40 feet | TV, Photo (Orbital Mission) |
| | soil properties | | 2-12 inches | TV, Photo (Entry & Impact Mission) |
| | volcanic activity | Visible | 5-20 feet | TV, Photo (Orbital Mission) |
| | faults | | 1-6 inches | TV, Photo (Entry & Impact Mission) |
| | lineaments | | | |
| Surface Imaging | Site Survey | IR | 1 inch - 1 foot | TV, Photo |
| | topography | Visible | 1 inch - 1 foot | TV, Photo, D/E Tape Camera |
| | color | UV | 1 inch - 1 foot | TV, Photo |
| | temperature | | | |
| | soil structure | | | |
| | Petrography | Visible | Microscopic | TV, Photo, D/E Tape Microscopes |
| | rock texture | | | |
| | bulk mineralogy | X-ray | Microscopic | Photo, others uncertain |
| | Mineralogy | | | |
| | Life Detection | Visible | Microscopic | TV, Photo, D/E Tape Microscopes |
| Meteorology | Atmospheric composition | µwave | 10 km | Radiometer |
| | Temperatures | IR | 10 km | Radiometer |
| | Cloud Movement | | 2000 feet | TV |
| | Cloud Structure | Visible | 1000 feet | TV, D/E Tape Camera |

(5) D/E Tape Camera.--Similar to a photographic system with the film replaced by a reusable dielectric tape employing electronic record/reproduce techniques.

## Imaging sensors

Thermal detectors and photodetectors.--The basic function of the imaging sensor is to convert electromagnetic radiation emitted or reflected by the scene into usable electrical signals at the sensor output. Devices capable of performing this function can be grouped generally as antennas, thermal detectors and photodetectors. The latter two groups can be further subdivided in terms of detection processes as shown in Table II. A brief description of each of these detection processes is provided in the following list.

(a) Bolometric.--Changes in temperature of the responsive element induced by the incident radiation cause changes in the electrical conductivity of the element. These changes in conductivity are monitored electrically.

(b) Thermovoltaic.--The temperature of a junction of dissimilar metals fluctuates because of changes in the level of incident radiation, giving rise to a fluctuating voltage generated by the junction.

(c) Thermopneumatic.--The radiation incident on a gas in a chamber increases the temperature (and therefore the pressure) of the gas causing the chamber to expand. This expansion may be detected optically.

(d) Photovoltaic.--A change in the number of photons incident on a p-n junction causes fluctuations in the voltage generated at the junction.

(e) Photoconductive.--A change in the number of incident photons causes a fluctuation in the number of free charge carriers in the semiconductive material. The electrical conductivity of the responsive element is proportional to the number of photons. This change in conductivity is monitored electrically.

## TABLE II
## THERMAL AND PHOTO DETECTION PROCESSES

| Group | Detection Process | Examples |
|---|---|---|
| Thermal Detectors | Bolometric | Thermistors |
| | Thermovoltaic | Thermocouples<br>Thermopiles |
| | Thermopneumatic | Golay cells |
| Photodetectors | Photovoltaic | Gallium Arsenide Cells<br>Indium Arsenide Cells<br>(Photographic Exposure Meter) |
| | Photoconductive | Television tubes,<br>Photodiodes,<br>Phototransistors |
| | Photoelectromagnetic | Indium Antimony Cells<br>Mercury Tellurium Cells |
| | Photoemissive | Image Dissector |
| | Photochemical | Film |

(f) Photoelectromagnetic. --Photons absorbed at the sensor surface generate charge carriers which diffuse into the bulk and are separated en route by a magnetic field. This separation of charge produces an output voltage which fluctuates according to fluctuations in the number of incident photons.

(g) Photoemissive. --Incident photons impart sufficient energy to electrons on the photoemissive surface to liberate them from the material. The freed electrons may be swept to an anode, and the resulting current can be monitored to determine fluctuations in the number of incident photons.

(h) Photochemical. --Photons incident upon an emulsion containing sensitive compound crystals (e.g., silver halides) partially reduces the compound crystals to small grains of the basic element (e.g., silver). The "latent image" produced is made visible by chemical development which completes the reduction of crystals affected by the exposure to photons. Resulting density of the developed film provides a measure of the number of incident photons over a given area.

The number of available devices in each process category is quite large. To attempt to catalog the devices and their parameters is beyond the scope of this study. It is sufficient to state that the experiment designer has a large number of detectors from which to choose for optimization of his experiment.

At this point it seems worthwhile to present illustrative diagrams of several types of imaging sensors to indicate the basic features of image sensor systems.

Microwave radiometer. --A Dicke-type microwave radiometer is illustrated in Figure 2. In operation, target noise power is received by the antenna and compared by the electronically switched circulator to the noise power generated by the reference source. The circulator output is amplified in the IF section of the receiver and synchronously detected, to determine the difference between receiver noise plus reference noise and receiver noise plus target noise. The detector output is integrated over many periods and subsequently passed to the recording system. For meteorological missions the system would be designed to operate at a wavelength of 13.5 millimeters ($\sim$ 24 GHz), the center of the main water vapor absorption line. For surface mapping missions the operating wavelength would probably be near 19 millimeters ($\sim$ 16 GHz), the wing of the water vapor absorption line, thus minimizing atmospheric interference.

Side-looking radar. --A simplified functional diagram of the side-looking radar appears in Figure 3. In operation the system transmits a pulse which is reflected by each resolution element of the target. The return signal is received and processed by an optical or electronic correlation device and the resulting image is placed in storage. In practice the system will probably use pulse compression techniques to minimize peak power requirements and improve range resolution capability. High angular resolution information is available by detecting the dopper shift on the pulse return from each elemental target; an on-axis target yields zero doppler. The correlator effectively sorts and sums the many returns from each elemental target as long as it remains within the unambiguous field-of-view of the system.

13
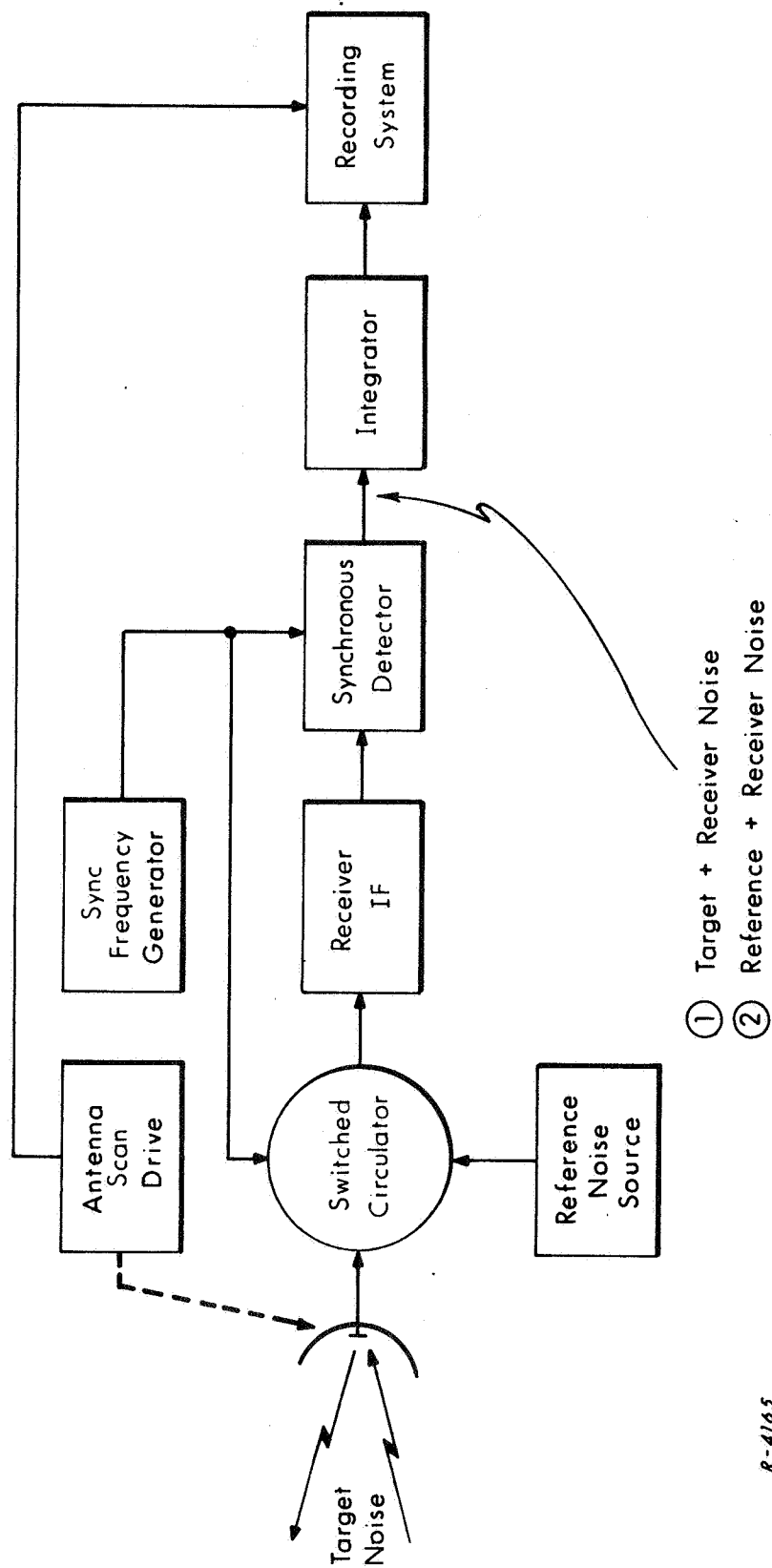
Figure 2  Dicke-Type Microwave Radiometer

① Target + Receiver Noise
② Reference + Receiver Noise

R-4165
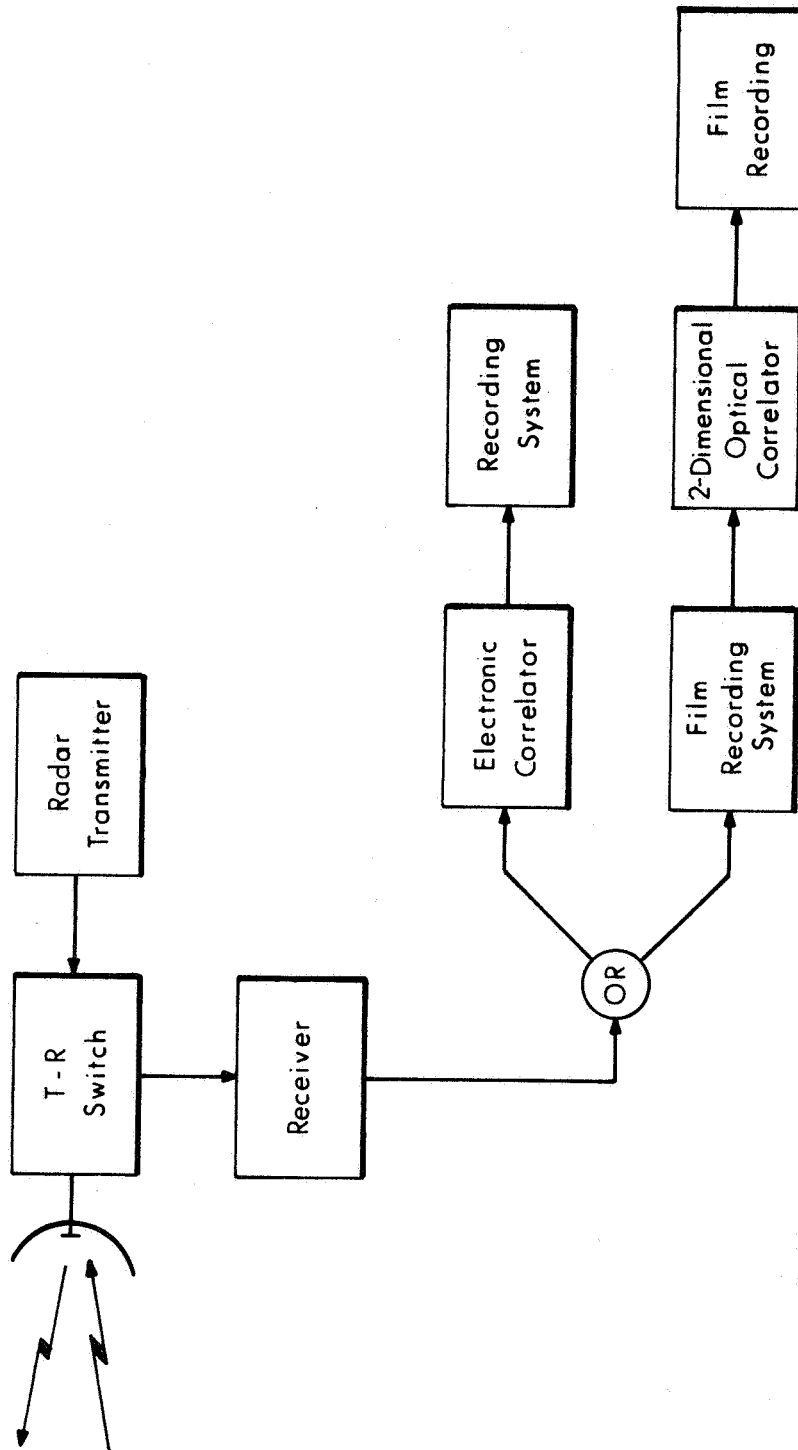
Figure 3  Side-Looking Radar

R-4164

If optical correlation were used, it is unlikely that the correlator would actually be flown; raw data would be transmitted from the vehicle to Earth. The advantage of the optical correlator is the improvement in resolution which can be realized over the electronic system. However, this advantage may be degraded if the intermediate transmission system is utilized.

While radar techniques may be desired for high resolution mapping of cloud-covered planets, their weight may make such a mission undesirable.

Infrared radiometer. --One type of infrared radiometer is illustrated in Figure 4. Note the similarity between this radiometer and the microwave radiometer of Figure 2. The operating principles are identical; the mechanization is very similar with all components in front of the radiation detector replaced with IR wavelength equivalents. The operating wavelengths of the IR radiometer will depend entirely on the experiment objective and atmospheric composition of the observed planet. Table III lists a few of the observation bands which may be of interest.

TABLE III
SOME IR WAVELENGTHS OF POSSIBLE INTEREST

| Wavelength | Observable |
|---|---|
| 6.0 microns | Atmospheric water vapor |
| 4.3 microns | $CO_2$ (e.g., Venus meteorology) |
| 3.5 microns | CH chain (Sinton bands) |
| 15.0 microns | $CO_2$ |
| ? | $NH_3$ (e.g., Jupiter meteorology) |

Television (TV). --Basic television sensor systems are well known and therefore no block diagram illustration is provided. The basic elements of the system are the optics, a shutter for exposure control and the imaging sensor. A variety of imaging sensors are available including solid-state mosaics, image dissectors, vidicons, image orthicons, return-beam vidicons, SEC vidicons and intensifier vidicons. These devices exhibit a wide variety of performance characteristics, some of which are presented in Table IV. The values shown in the table are dependent upon actual experiment design requirements and therefore are only intended to indicate the wide range of performance characteristics available to the designer.

16
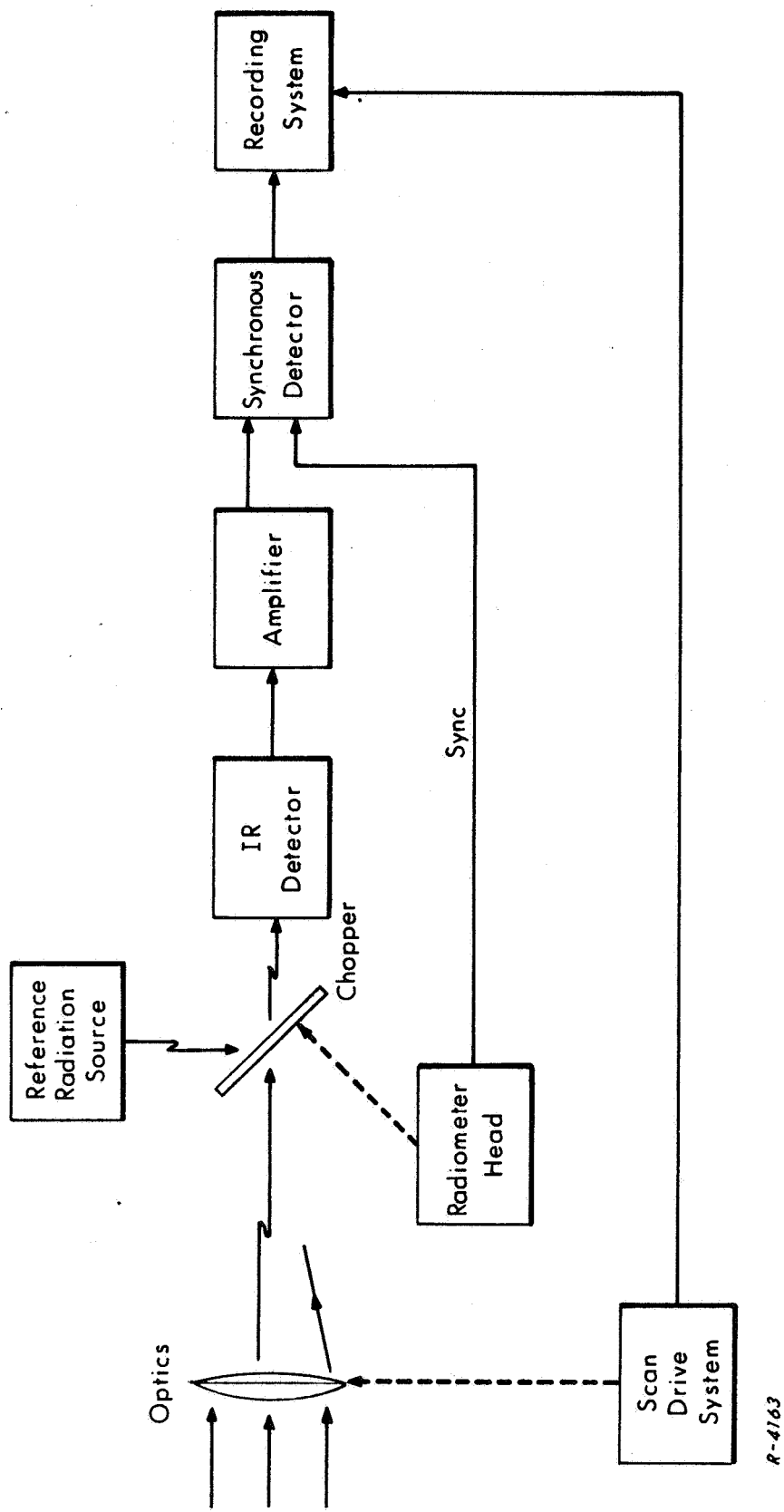
Figure 4  Infrared AC Radiometer

R-4163

## TABLE IV
## TELEVISION SENSOR CHARACTERISTICS

| Characteristic<br>Sensor | Sensitivity*<br>(foot-candles) | Line<br>Resolution | Dynamic<br>Range | Ruggedized | Weight<br>(pounds) |
|---|---|---|---|---|---|
| 1/2" Vidicon | 0.2 | 400 | 100:1 | Yes | 1 |
| 1" Vidicon | 0.1 | 1000 | 100:1 | Yes | 3 |
| 1-1/2" Vidicon | 0.04 | 1100 | 100:1 | No | 3 |
| 2" Return Beam Vidicon | 0.02 | 4000 | 100:1 | No | 10 |
| 4-1/2" Return Beam Vidicon | 0.007 | 7000 | 100:1 | No | ~30 |
| SEC Vidicon | $7 \times 10^{-4}$ | 300 | 60:1 | No | 1/2 |
| 2" Image Orthicon | $3 \times 10^{-5}$ | 100-800<br>(function of<br>light level) | 15:1 | Yes | 5 |
| Intensifier Vidicon | $10^{-5}$ | 400 | 1000:1 | No | 1/2 |
| Solid-State Mosaic Development | 0.1 | 128 | >100:1 | Yes | ? |
| Image Dissector | $10^{-4}$ | 3000 | >100:1 | ? | 16 |

* Sensitivity at S/N = 10 and exposure time equal to 1/30 seconds.

Photographic sensor system. --The basic principles of operation of photographic cameras are well known and do not require description here. The greatest disadvantages of photographic systems are the chemical development requirements, the film volume required for large quantities of pictures and the radiation sensitivity of exposed film. These disadvantages have been overcome with the development of dielectric tape camera systems. However, such parameters as sensitivity and resolution are inferior to standard photographic cameras. The weight of the dielectric tape camera system exceeds 80 pounds.

## Fidelity criteria

Qualitative criteria. --The purpose of this section is to point out a few of the problems one encounters in the attempt to define a fidelity criterion for imaging systems. A review of available literature indicates that three basic picture characteristics determine image quality.

(1) The tone or color contrast between an object and its background.

(2) Image sharpness.

(3) Resolution capability.

In order for an imaging system to render tones accurately it must have a gray scale rendition suitable in dynamic range and linearity to the image requirements of the mission. The impression of image "sharpness," while it is to some extent a subjective one, depends primarily upon the sharp rendition of abrupt boundaries between areas of different tone. The retention of these sharp edge contrasts requires good spatial step response in the system.

In the observation of geological features one may encounter thin line details at virtually the limit of resolution. The output of a scanner sweeping such a picture will contain impulses. Clearly, the ability of a sensor system to convey in an identifiable manner the finest pictorial details, depends upon its impulse response.

In a linear electrical network the step response s(t) is related to the impulse response h(t) by

$$s(t) = \int_0^t h(\tau)d\tau \qquad (2)$$

Good impulse response and good step response may be seen from Eq. (2) to be not incompatible requirements. A clean impulse response will have a sharp central peak and little ripple after this narrow peak. Its integral will be a sharp-edged step with little ripple on top. In an analogous manner image sharpness and resolution characteristics do tend to be mutually attainable objectives.

In some scenes, either geological or meteorological, adjacent areas may be noted which appear to have approximately equal average tone. However, there may be a distinct demarcation between these areas on the basis of texture. For example, large old-growth timber closely packed

differs in texture from similar species timber of new growth and medium density, yet the average tone for the areas is quite close. Even though individual trees are below the system resolution limit, a system with good midrange sinewave response may provide adequate texture distinction.

The obvious question of how to relate the response characteristics described in the preceding discussion to analytical fidelity criteria is discussed in the following section.

Quantitative criteria

General considerations.--A fidelity criterion is a rule to be used for the ranking of data specimens according to their quality or utility. As previously noted, the measure of data fidelity is usually a function of the error between data obtained and its true value, considered over some observation interval:

$$\Phi \equiv F\left[\hat{x}(t) - x(t)\right] = F\left[e(t)\right]$$

where the fidelity measure $\Phi$ is a scalar quantity determined by the functional operation $F[\ ]$, which denotes the fidelity criterion.

The choice of an effective fidelity criterion depends upon the use to which the data are to be put and the methods and techniques used in their interpretation. Because of the subjective nature of these considerations, choosing a criterion which discriminates accurately and consistently between data of greater and lesser quality is generally a very difficult task.

It is convenient to divide fidelity criteria into two classes. One class is composed of those most sensitive to long-term or average error and the other those most sensitive to short-term or localized errors. In general, fidelity criteria should be chosen to correspond as closely as possible to desired data information.

Short-term fidelity criteria should be selected in the event that short-term information is desired from the data, so an accurate and sensitive assessment of system performance can be made. A similar selection of long-term fidelity criteria should be made if long-term information is desired.

Average fidelity criteria. --The following average fidelity criteria are defined in terms of a general time waveform of error, e(t), which is assumed to be 0 for all t < 0. If it is assumed that the error process is ergodic, error measurements derived by time averaging techniques can

be shown to converge with probability 1 to the corresponding quantities obtained by statistical averaging techniques. Since the ensemble of error waveforms required for implementation of statistical techniques is often physically unrealizable, time averaging techniques based upon a single stationary error waveform become an acceptable and entirely equivalent alternative.

Time-average error (TAE). --Based upon an averaging period, T, TAE is mathematically defined as

$$TAE = \frac{1}{T} \int_0^T e(t)dt$$

As T is allowed to become arbitrarily large, time-average error converges with probability 1 to the statistical mean error, $\mu_e$.

Time-average absolute error (TAAE). --Based upon T, TAAE is mathematically defined as

$$TAAE = \frac{1}{T} \int_0^T |e(t)|dt$$

Time-average squared error (TASE). --Based upon T, TASE is defined as

$$TASE = \frac{1}{T} \int_0^T e^2(t)dt$$

As T is allowed to become arbitrarily large, time-averaged squared error converges to the statistical mean squared error, with probability 1.

Root time-average squared error (RTASE). --This is the square root of the time-average squared error. As the averaging period T is allowed to become arbitrarily large, this quantity converges with probability 1 to the statistical root mean squared (rms) error.

Time-average error variance (TAEV).--Based upon $T$, TAEV is mathematically defined as

$$TAEV = \frac{1}{T} \int_0^T \left( e(t) - \frac{1}{T} \int_0^T e(a)da \right)^2 dt$$

$$= \frac{1}{T} \int_0^T e^2(t)dt - \left( \frac{1}{T} \int_0^T e(t)dt \right)^2 = (TASE) - (TAE)^2$$

As $T$ is allowed to become arbitrarily large, this quantity converges with probability 1 to the statistical error variance, $\sigma_e^2$.

Time-average error standard deviation (TAESD).--This is the square root of the TAEV. As $T$ is allowed to become arbitrarily large, this quantity converges with probability 1 to the statistical error standard deviation, $\sigma_e$.

Short-term fidelity criteria.--All average fidelity criteria described in the previous paragraph also can be considered short-term fidelity criteria when $T$ becomes short with respect to the minimum time interval required for appreciable variation in the error waveform. The dividing line between average and short-term fidelity criteria is often quite indistinct.

Additional short-term fidelity criteria which are not derivable from the criteria of the preceding paragraph are the maximum absolute error and the peak absolute error.

Maximum absolute error (MAE).--Based upon an observation period, $T$, MAE is mathematically defined as

$$MAE = \underset{T}{MAX} ( |e(t)| )$$

where $\underset{T}{MAX}( )$ denotes the maximum value of ( ) during this observation period of length $T$.

Peak absolute error (PAE). --PAE for an observation period, T, is that value of $|e(t)|$ which is exceeded an arbitrarily small fraction of the time, $\epsilon$, i.e.,

$$p(|e(t)| > PAE) = \epsilon$$

For $\epsilon = 0$, MAE and PAE are equal.

Based on the qualitative discussion under the section Qualitative Criteria, page 19, which indicates a need for good step function and impulse responses in the system, it appears that the short-term fidelity criteria would be most applicable to imaging data systems. Furthermore, minimization of the peak absolute error (PAE) appears to be the more practical fidelity criterion for television and photographic type imaging sources. No conclusions are possible at the present time concerning fidelity criteria selection for radiometric and mapping radar imaging sources.

## Second order interpolation algorithm

Simulation of the algorithm. --The second order interpolator (SOI) attempts to fit a parabola to a sequence of data values such that the parabola (1) passes through the first data point; (2) passes no more than a certain prespecified distance away from each succeeding data point; and (3) continues in this fashion for as many data points as possible. The data points which are represented in this way are describable in terms of the approximating parabola's parameters, hopefully with a savings in information required for their description.

A geometric description of SOI operation is most convenient. The general equation of an approximating parabola is

$$\hat{y}_n = an^2 + bn + c, \qquad n = 0, 1, 2, \ldots \qquad (3)$$

where $\hat{y}_n$ is the value of the parabola corresponding to the $n^{th}$ sample time. Since any parabola is completely described as the three parameters a, b, and c, any given parabola can be described by an appropriate point $(a_0, b_0, c_0)$ in a three-dimensional space with coordinate axes a, b, and c (see Figure 5). Moreover, given any three data values, $y_0$, $y_1$, and $y_2$, it can be shown that all parabolas satisfying the first two conditions imposed on the SOI above can be represented by the interior points of a two-dimensional parallelogram in a-b-c space (see Figure 6). Given data values $y_0$, $y_1$, $y_2$, and $y_3$, it can be shown that all parabolas satisfying the first

23

$$\hat{y}_n = a_0 n^2 + b_0 n + c_0$$
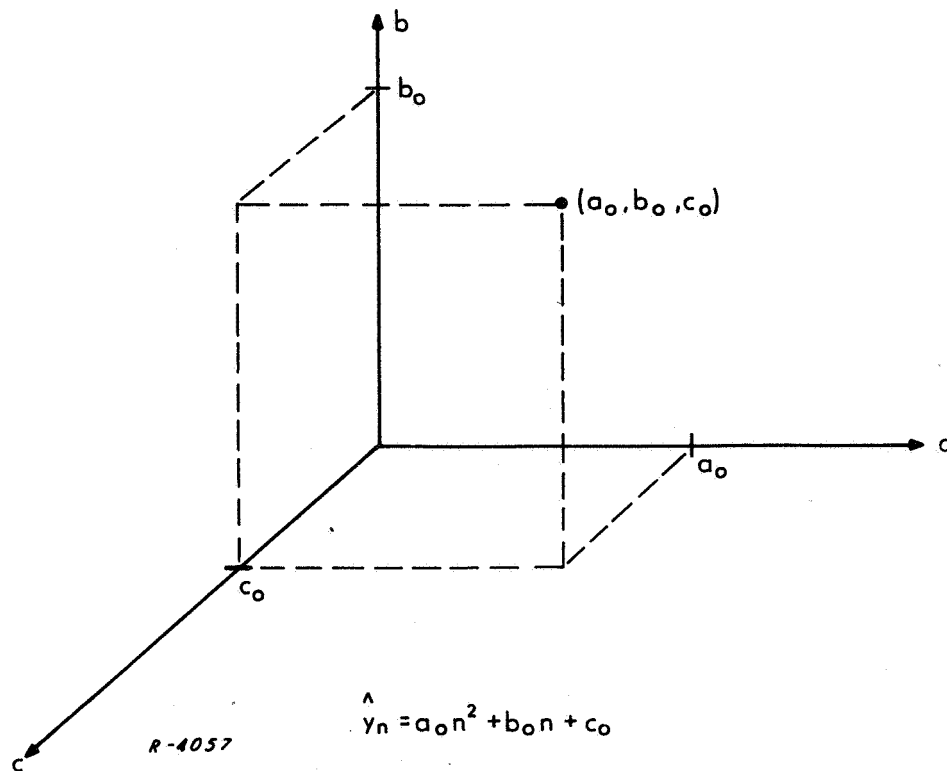
Figure 5  Representation of a Parabola as a Point
in Three-Dimensional Space
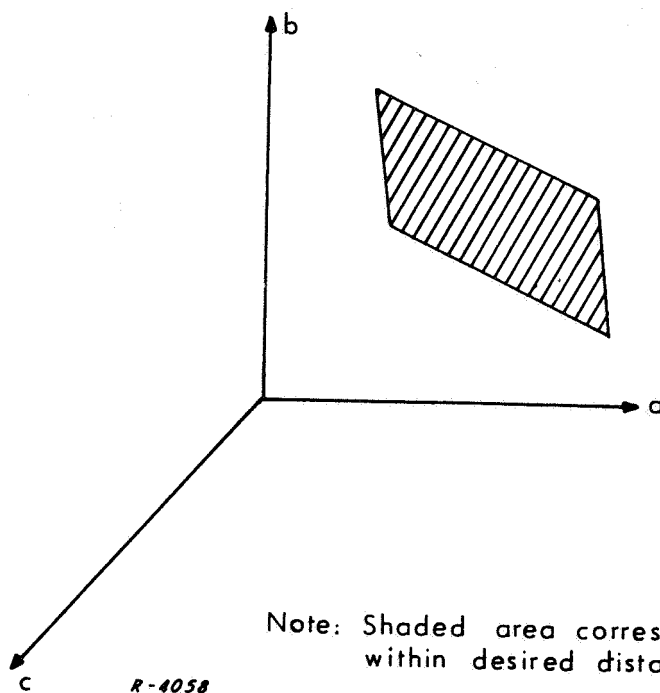


Note: Shaded area corresponds to all parabolas within desired distance of $y_0$, $y_1$, and $y_2$

Figure 6  Geometric Representation of all Parabolas Satisfying
Maximum Error Criterion with Respect to Three
Data Points

two conditions imposed on the SOI can be represented by the interior points of the intersection of the two coplanar parallelograms corresponding to $y_0$, $y_1$, $y_2$ and $y_0$, $y_1$, $y_3$ (see Figure 7). This procedure may be extended, one data value at a time, until the intersection of these coplanar parallelograms becomes a null space, indicating it is no longer possible to represent the given data values within the imposed maximum error constraint. Any point in the previous intersection, however, corresponds to a parabola which describes all samples but the last within the required maximum error constraint. New parabolas may be initiated in this manner until all data points have been represented.
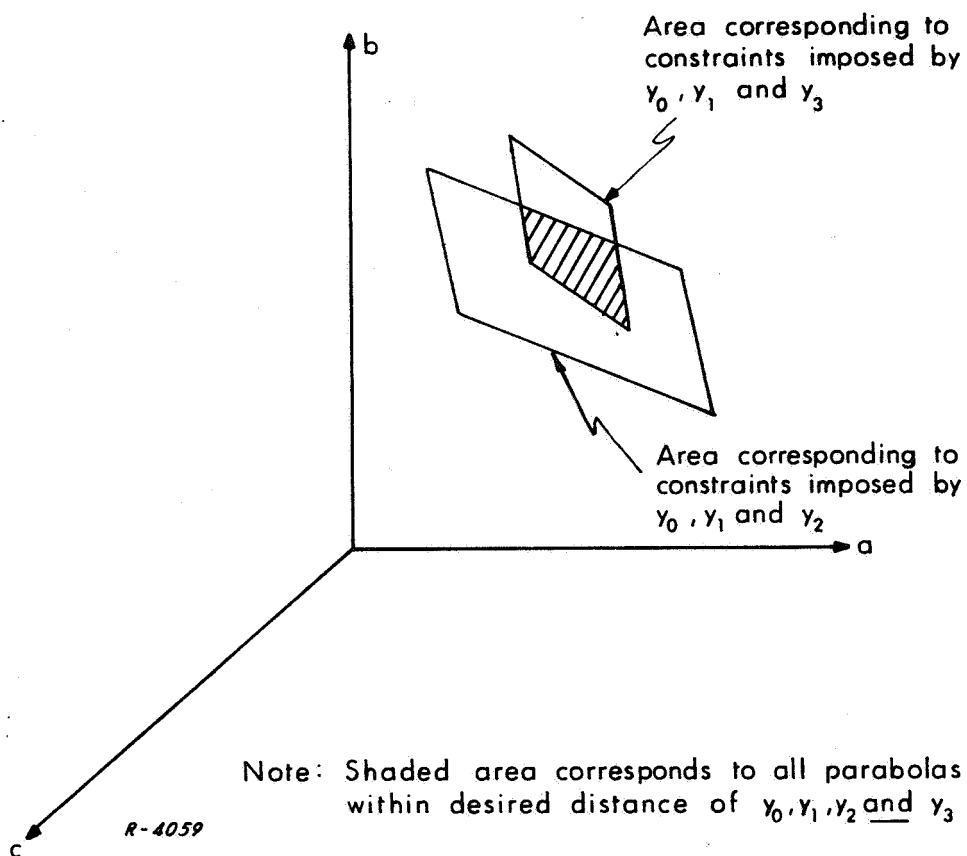


Figure 7    Geometric Representation of all Parabolas Satisfying Maximum Error Criterion with Respect to Four Data Points

The operation of second order interpolation necessarily produces line-segment compression ratios[*] which are at least as high as the compression ratios resulting from first order interpolation. This is so because of the fact that second order interpolation reduces to first order interpolation if the parameter "a" is constrained to zero, making FOI a special case of SOI. No general statement can be made, a priori, about the relative efficiency (in terms of bandwidth compression capability) of FOI and SOI compression algorithms, since bandwidth compression achieved depends not only on line-segment compression ratio achieved, but also on the amount of information required to characterize each line segment. SOI requires more information per line-segment than does FOI. Consequently, in some instances FOI may be more efficient in terms of bandwidth compression than SOI, simply because the superior line-segment compression ratio achieved by SOI is not sufficient to offset the increased amount of data needed to specify a line segment relative to FOI.

Line-segment compression ratios obtained by compressing a portion of a high-resolution reconnaissance photo with various algorithms, including SOI[**], are presented in Figure 8. The original picture (after scanning) is shown in Figure 9 and the compressed and reconstructed versions using FOIDIS and SOIDIS are shown in Figures 10 and 11, respectively.

The basic experimental procedure for compressed picture simulation is shown in block diagram form in Figure 12. Several experimental parameters are noted on the diagram. The "algorithm simulation only" path led to the production of Figures 9-11. Figures 14-16 were generated by following the "noisy channel simulation" path; this effort is discussed under the section Simulation of Effects of Channel Errors Upon Compressed Pictures, page 32.

---

[*] Line-segment compression ratio is defined as follows:

$$N_{\text{line-segment}} \triangleq \frac{\text{\# of sample points processed}}{\text{\# of line segments required}}$$

where the term "line segment" is understood to signify a segment of a parabola when reference is to the SOI.

[**] In accordance with established convention, the simulated second order interpolation algorithm was named, "Second Order Interpolation, Disjoined" with the acronym SOIDIS, because of the disjoined nature of successive approximating line segments.
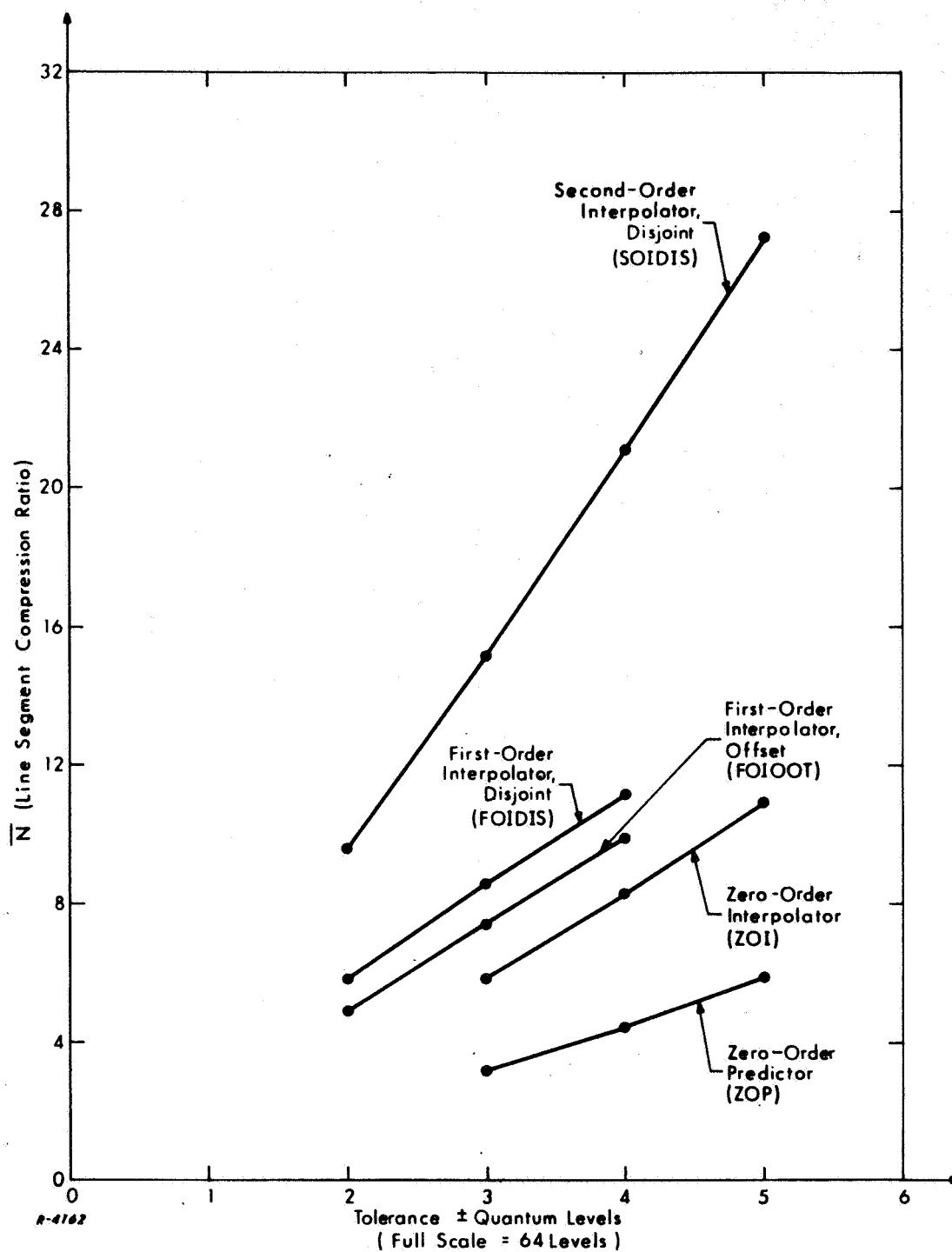
Figure 8 Performance of Compression Algorithms on Segment
of High Resolution Reconnaissance Photo

Fig. 9    Original Picture (Scanned and Reconstructed)

28
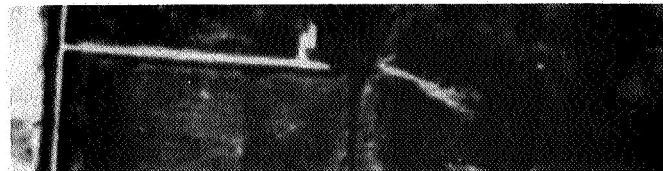
Tolerance = 2, $\overline{N}$ = 5.78



Tolerance = 3, $\overline{N}$ = 8.48



Tolerance = 4, $\overline{N}$ = 11.08

Fig. 10    Compression Results (FOIDIS)

Tolerance = 2, $\overline{N}$ = 9.64



Tolerance = 3, $\overline{N}$ = 15.15



Tolerance = 4, $\overline{N}$ = 21.06



Tolerance = 5, $\overline{N}$ = 27.2

Fig. 11    Compression Results (SOIDIS)
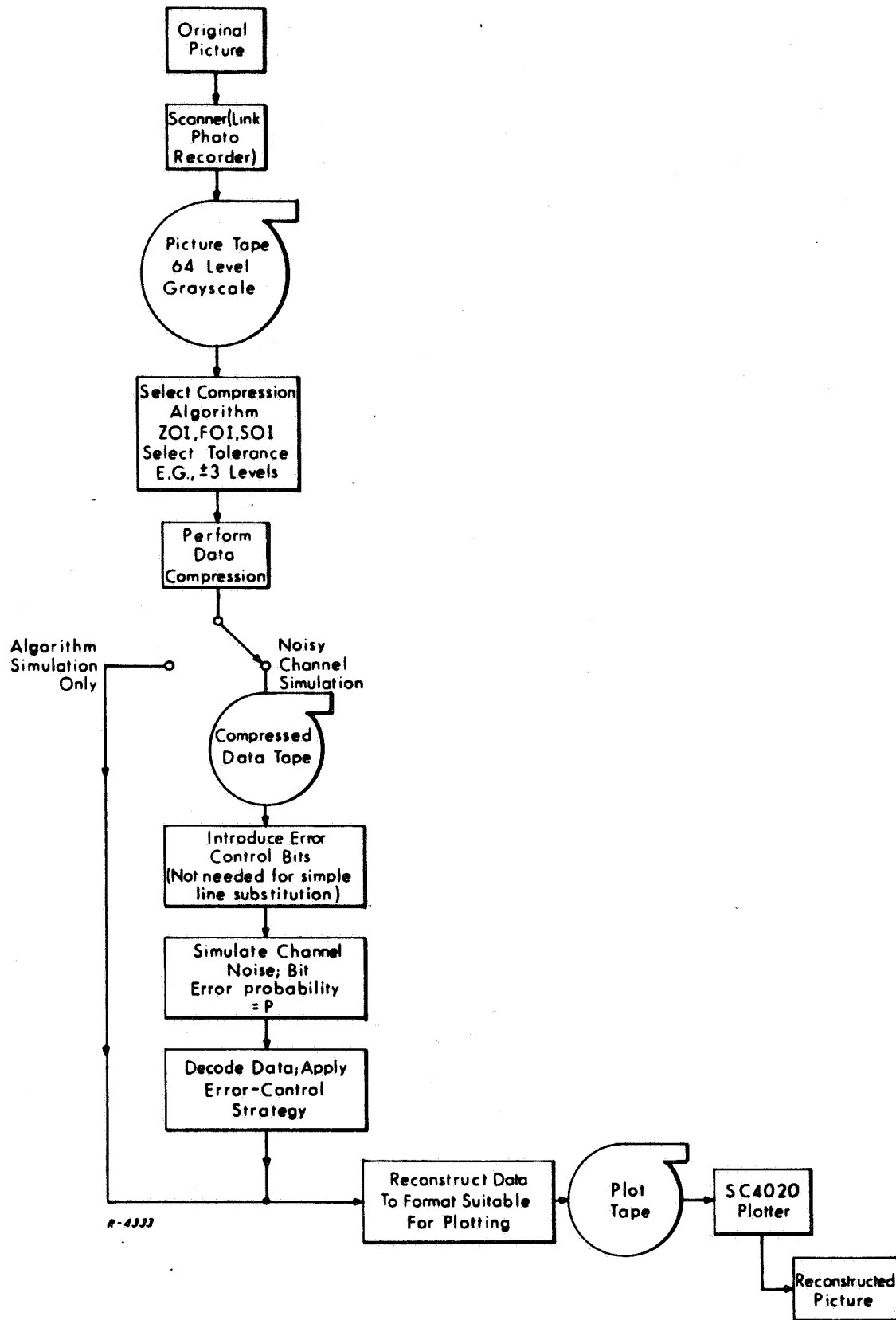
Figure 12  Block Diagram of Basic Algorithm Simulation Approach

A note about picture quality must be injected at this point. Figures 9-11 were done using an SC4020 plotter, chosen on account of convenience and accessibility. The SC4020 was originally designed for plotting charts and graphs and does not produce high-quality rendition of continuous-tone picture material. Occasional data dropouts, occurring when the plotter fails to properly read portions of the plot tape, produce spurious horizontal white line segments across the picture. Arrangements have been made for access to other plotting equipment, and pictures of good technical quality will be available for the final report.

Compressed data quality for SOIDIS appears as good or better than that resulting from use of FOIDIS at an equivalent tolerance. SOIDIS is seen to produce line-segment compression ratios approximately twice as great as FOIDIS. If the four parameters of SOIDIS (a, b, c, and runlength) can be efficiently encoded to require less than twice as many bits as the three of FOIDIS (b, c, and runlength) a net gain in bandwidth compression will be realized.

## Future effort in algorithm simulation

Further work on the SOI will include items such as:

(1) determination of the relation between line segment and bandwidth compression ratios, and

(2) modification of the SOIDIS algorithm to allow SOIOOT[*] operation with subsequent determination of compression efficiency.

## Simulation of effects of channel errors upon compressed pictures

General. --This phase of the study concerns itself with the sensitivity of continuous-tone monochrome picture data to channel errors when it is transmitted in digitally coded form over a binary symmetric channel.

Figure 13 shows diagrammatically the basic steps comprising a picture error sensitivity simulation. It is to be expected that the different algorithms will produce characteristically different binary picture data

---

[*]Constraints upon SOIOOT line segments are analogous to those describing FOIOOT: each segment must have a fixed y-displacement (+ or -) with respect to the former. Therefore, only one bit is required to locate the beginning of a segment with respect to the end point of the preceding segment. SOIOOT consequently requires fewer bits to specify a segment than does SOIDIS.

32

Select Pictorial Subject → Select Compression Algorithm; Generate Set of Significant Samples → Express Picture Data in Binary Form; Add Error Control IF Desired → Simulate Errors Binary Symmetric Channel, Transition Probability = p → Decode Data; Reconstruct and Plot Replica of Original Picture
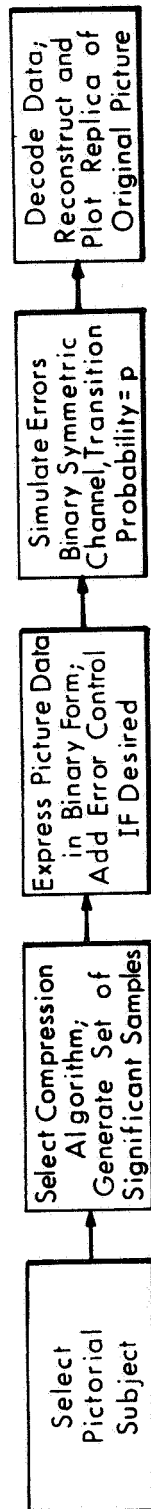
R-4174

Figure 13   Basic Steps in Picture Data Error Sensitivity Simulation

33

which are differently affected by channel noise. The channel noise model
to be used throughout this effort will be the binary symmetric model in
which bits suffer independent and random alteration in accordance with a
transition probability p. The details of the error simulation are described
in Appendix A.

Zero order interpolator. --Zero order interpolation effectively
replaces the data waveform with a staircase approximation shown below.
Each horizontal line segment (significant data sample) replaces several
samples; the compressed data lies within a tolerance width $\epsilon$ ($\pm \epsilon / 2$)
about the original data. Each significant sample is specified in terms of
two numbers: the sample value and its location (or alternately a runlength
from the preceding sample).

The zero order interpolator was chosen as the first algorithm to
employ for this simulation. It is considerably easier to design software
for it than for the first order interpolators, and it has been widely used
in the past for the compression of pictorial material. A high resolution
aerial reconnaissance photograph was selected and scanned to form 450
lines, each consisting of 450 samples from a 64-level gray scale quanti-
zation
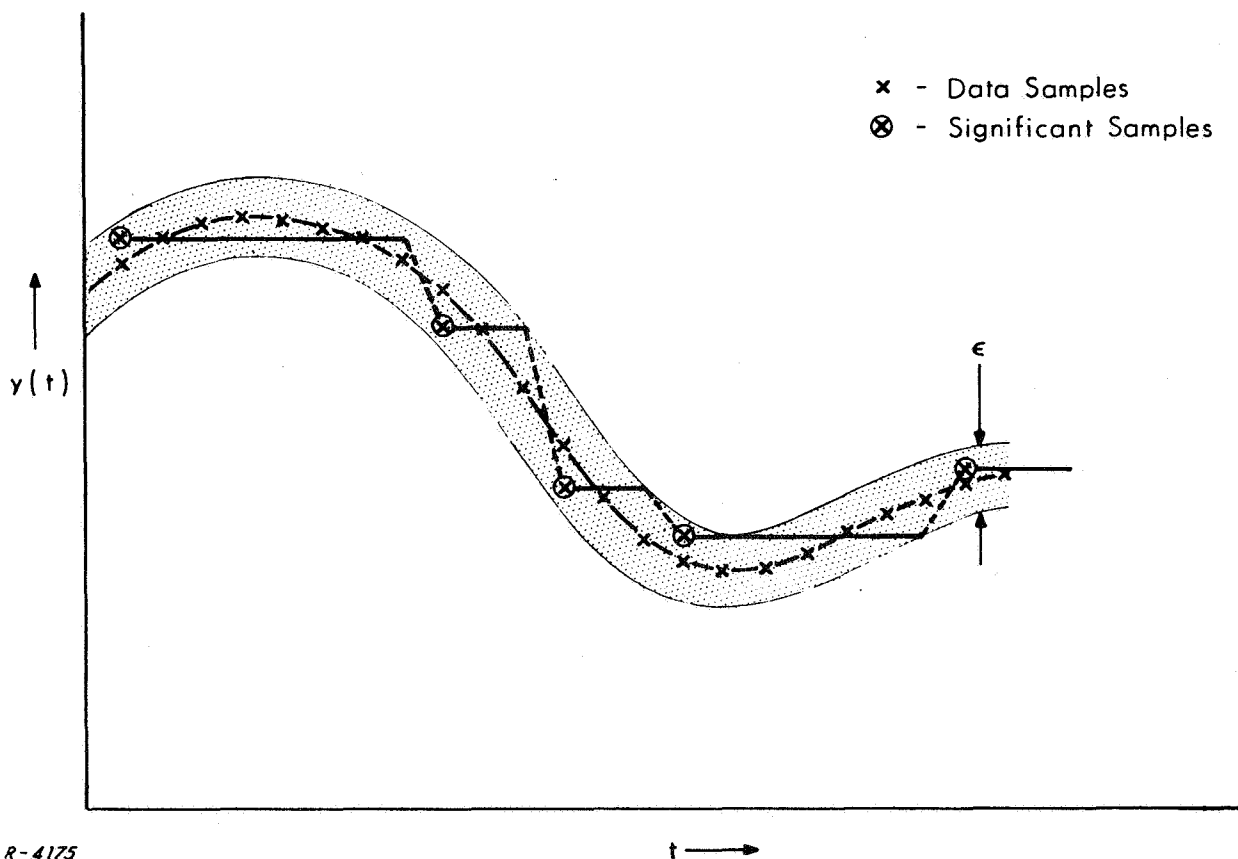
After the generation of the significant sample representation by application of the algorithm, the data was formated as follows. Each significant sample was represented by 11 bits; six bits constitute the gray scale value by straight binary mapping, and five bits represent the runlength, i.e., the spacing between this significant sample and the preceding one. The degradation in algorithm compression capability caused by this limitation upon maximum runlength (32) is insignificant. Note, however, that the reduction in sample addressing data by using five bits per sample rather than the ten required to indicate one location out of 450 is quite substantial. It is further considered that the lines of picture data are separated by a line synchronization sequence which is completely immune to errors.

These assumptions about data format closely approximate what would be done in any realistic picture transmission scheme. For the minimization of bandwidth, a runlength data point addressing scheme would almost certainly be chosen in preference to a straight location addressing scheme. Also, the use of any reasonable binary sequence for line sync (one with at least a few bits more than the eleven required to specify a data point) will lead to a situation where the probability of line sync failure is several orders of magnitude lower than the probability of an error within the data bits describing a picture line. For data error control schemes which are not of high complexity those error rates so large as to make line sync failure a significant occurrence will most likely result in total degradation of the picture information, and thus are not of major interest.

Results. --The result of compressing the picture to a tolerance of $\pm 3$ quantum levels of true value are shown in Figure 14. Figure 15 shows the degrading effects of a $10^{-3}$ bit error probability. Errors affecting a five-bit runlength word are, for this error probability, present in approximately one-third of the scan lines. This sort of error introduces a shift in the line which proceeds from the point of error to the end of the line. Ragged edges at the extreme right of the picture are indicative of the fact that such lines have a length that differs from 450 cells. Ragged edges also appear at intensity boundaries within the picture. (Due to poor definition of the SC4020 plotter when this picture was processed, some of the degradation is masked by picture fuzziness.) The effects created by run word errors are definitely seen to be more serious degrading factors than the "off-color" bars produced when an intensity level error raises or lowers one of the horizontal line segments of the staircase. Errors in intensity value affect only the portion of a line represented by the sample whose intensity has been changed; runlength errors affect all succeeding samples on the line.

ZOI, Tolerance = 3, $\overline{N}$ = 5.77

Fig. 14    Compressed Reconnaissance Photo
Noise Free

ZOI, Tolerance = 3, $\overline{N}$ = 5.77

Fig. 15    Compressed Reconnaissance Photo
Noisy Channel, p = $10^{-3}$

In view of the greater significance of run word errors, the following simple error control scheme was felt to have potential for improving data quality over a noisy channel. Consider that one has an accumulator, reset to zero after each line sync, which adds the total of all the runlength words since the beginning of the line. Upon the arrival of line sync, the contents of the accumulator are read. If the accumulator contents total 450, the line is considered to be satisfactory and may be used in the construction of the picture. If the total does not equal 450, the line certainly has an error in it, and is declared to be unusable. The preceding line, which has meanwhile been held in storage, is used to fill in for the destroyed line. On account of the high line-to-line correlation of picture information, this line substitution procedure should result in acceptable picture quality, even when a sizable portion of lines are substitutes.

Figure 16 shows the consequences of a $10^{-3}$ error probability when this line substitution technique is employed. The improvement has been achieved with little complexity: a special memory large enough to store a line (most certainly < 2000 bits in this example), and an accumulator capable of totaling up to a 10-bit number. Observe that this approach will properly detect all lines having a single runlength word in error. When there are two words in error on a line, this will seldom fail to be detected (about 10% of the time). There is always a previous line available for substitution use (the same line will need to be used two or more times when consecutive lines fail) once the first line has been received with runlengths that total up properly. If the picture should happen to begin with one or more unusable lines, then the beginning of the picture is totally suppressed.

The next step to be pursued is the repetition using a first order algorithm of what has been done using the zero order interpolator. Picture evaluation must then be performed. The primary purpose of this evaluation will be to judge how effective line substitution is as an error control technique. It is additionally hoped that this picture evaluation may help in the generation of ideas for the creation of error control techniques one step in complexity beyond line substitution.

## Simulation test of A/D mapping technique

General. --This portion of the contract effort pertains to a novel A/D mapping technique by which $\ell$-bit sequences are assigned to each of the $2^{\ell}$ levels comprising the pictorial gray scale in such a way that a single bit error in any $\ell$-bit sequence translates into an error of at least $2^{\ell-2}$ levels (1/4 of full gray scale). The average value of displacement due to a single error when all levels are equiprobable is $2^{\ell-1}$ quanta (1/2 of full gray scale). The method of construction and derivation of the properties of this mapping are contained in Appendix B.

38

ZOI, Tolerance = 3, $\overline{N}$ = 5.77

Fig. 16    Compressed Reconnaissance Photo
Noisy Channel, p = $10^{-3}$
Line Substitution Error Control

The intent of the simulation described below was to test out the hypothesis that for data containing a reasonable degree of sample-to-sample correlation the use of this mapping technique will lead to a condition where occasional bit errors will produce "wild" data points. These wild points may be removed from the data by inspection, and the gaps filled in by interpolation between good points.

The simulation model. --The simulated data used for the test was quantized to 16 levels. It is generated by a Markov process: any given sample is dependent only upon the immediately preceding sample. A brief description of the generation of the pseudorandom data sequences is presented below.

The sample value $Y_i$ is a function of two generated random numbers. $U_i$ and $V_i$ are statistically independent and uniformly distributed over $0 \leq U_i, V_i < 1$. The new variable $Z_1$ defined by

$$Z_i = (U_i + V_i)/2 \tag{1}$$

has a triangular probability density function over the same unit line, with a peak at 0.5. $Y_i$ is determined by the following relationship, with A an odd integer:

$$Y_i = \text{int}(AZ_i) - \frac{A-1}{2} + Y_{i-1}$$

$$= Q_i + Y_{i-1} \tag{2}$$

The functional notation int denotes "The integer part of," e.g., int (0.3) = 0, int (7.75) = 7, int (3.000) = 3. The random variable $Q_i$ has zero mean and assumes integer values (including zero) ranging from $-(A-1)/2$ to $+(A-1)/2$. Table V illustrates the resulting tapered symmetrical probability distribution function of $Q_i$.

At this point it is clear how $Y_i$ is generated from the previous value $Y_{i-1}$ as long as the $Y_i$ may assume any integer values, positive or negative. In order to simulate 16 level quantized data, Y values must all be in the range $0 \leq Y_i \leq 15$. Therefore, any time the value of $Q_i$ generated produces a $Y_i$ falling out of the range, that value is ignored; new and independent values of $Q_i$ are generated until an acceptable $Y_i$ is produced.

It was assumed that a quantized level is represented by a 4-bit sequence in accordance with the mapping of Table VI. The production of error patterns corrupting the data is assumed to follow the binary symmetric channel model, i.e., each bit is independently affected by the transition probability p.

## TABLE V
## PROBABILITY DISTRIBUTION FUNCTION OF THE RANDOM VARIABLE $Q_i$

$$Pr \quad (Q_i \ = \ x)$$

| x | A = 11 | A = 15 |
|---|---|---|
| +7 | | .0089 |
| +6 | | .0266 |
| +5 | .0165 | .0445 |
| +4 | .0496 | .0622 |
| +3 | .0826 | .0800 |
| +2 | .1156 | .0977 |
| +1 | .1488 | .1156 |
| 0 | .1734 | .1290 |
| -1 | .1488 | .1156 |
| -2 | .1156 | .0977 |
| -3 | .0826 | .0800 |
| -4 | .0496 | .0622 |
| -5 | .0165 | .0445 |
| -6 | | .0266 |
| -7 | | .0089 |

TABLE VI
ACTUAL MAPPING USED FOR SIMULATION RUNS

$$
\begin{bmatrix}
15 \\
14 \\
13 \\
12 \\
11 \\
10 \\
9 \\
8 \\
7 \\
6 \\
5 \\
4 \\
3 \\
2 \\
1 \\
0
\end{bmatrix}
=
\begin{bmatrix}
1101 \\
1011 \\
0111 \\
0001 \\
1110 \\
1000 \\
0100 \\
0010 \\
1111 \\
1001 \\
0101 \\
0011 \\
1100 \\
1010 \\
0110 \\
0000
\end{bmatrix}
$$

The necessary software to perform the simulation was programmed in BASIC language for use on the GE 200 series Time-Sharing computer system, accessible through a teletype facility located at ADCOM, Cambridge. There are three fundamental parameters of the simulation controllable through the choice of initial program input parameters:

(1) the number of data points in the run;

(2) the range covered by the between-samples difference variable $Q_i$; and

(3) the bit transition probability.

The run of simulated data is generated twice. First the data are generated with the error-introducing operations of the program bypassed. The second run uses the entire program and produces the error-perturbed version of the same data. The wild-point discrimination and interpolation procedures are then manually applied to the data; the deviation of "corrected" data from the original unperturbed data may also be manually observed.

Analysis of results. --Two basic simulation runs were carried out each consisting of 500 data points. The first was for A = 11 (-5 $\leq Q_i \leq$ +5) and the second for A = 15 (-7 $\leq Q_i \leq$ +7). A bit error probability of 0.02 was simulated. This relatively high error probability, it must be realized, leads to a fairly conservative evaluation of the discrimination and interpolation technique. This is because wild points are most distinguishable when standing out between good data samples. However, when p = 0.02, 7.5% of the data samples contain errors, resulting in areas of error clustering in which discrimination will not work so well.

The following rules govern the discrimination and interpolation procedure:

(1) The data consists of point $Y_i$; i = 1, 2 ... 500.

(2) A decision is made upon these points one at a time in succession, beginning with $Y_1$.

(3) A data point $Y_i$ is judged to be good or bad upon the following criterion. If both $\Delta^+ \equiv |Y_{i+1} - Y_i|$ and $\Delta^- \equiv |Y_i - Y_{i-1}|$ are equal to or greater than some threshold value T, the point is considered to be "bad"; otherwise the point is "good." For the data point $Y_1$, we assume $\Delta^- > T$, and for $Y_{500}$ we assume $\Delta^+ > T$.

(4) If a point is bad, we use linear interpolation or extrapolation to fill it in as in rules (5) to (7) below. It is then considered to be "good" for future use. If $Y_i$ is good we go to $Y_{i+1}$ and apply our tests again.

(5) If $Y_i$ is bad, we attempt to find that $Y_j$ for the smallest $j > i$ which is "good" in accordance with rule (3). If there is such a point, and if $i \geq 2$, we correct $Y_i$ to the value $\hat{Y}_i$ by linear interpolation:

$$\hat{Y}_i = Y_{i-1} + (Y_j - Y_{i-1})/(j - i + 1)$$

The estimated value is rounded off to the nearest integer; if $\hat{Y}$ is halfway between two integers, it is rounded to the even integer.

(6) If it is necessary to correct the value of $Y_1$, we must extrapolate backwards using the first two good points to be found, $Y_j$ and $Y_{j+\alpha}$. $Y_1$ is corrected to $\hat{Y}_1$ by the linear extrapolation:

$$\hat{Y}_1 = Y_j - (j - 1) \cdot \left[ \frac{Y_{j+\alpha} - Y_j}{\alpha} \right]$$

(7) If $Y_i$ is near the end of the data run there may be no "good" $Y_j$ for $j > i$. Then $Y_i$ is corrected by extrapolation using $Y_{i-1}$ and $Y_{i-2}$.

$$\hat{Y}_i = Y_{i-1} + (Y_{i-1} - Y_{i-2}) \cdot$$

Before studying the actual characteristics of errors remaining in data runs which have been subjected to this interpolation procedure, we must have a standard comparison. The alternate approach to the use of a discrimination-correction technique in conjection with a mapping which causes single errors cause small quantum level distortions. The average displacement due to a single error is theoretically minimized by the use of straight binary mapping in which the levels 0 to 15 directly correspond to their four-bit binary representations, 0000 to 1111.

As shown in Appendix B, when straight binary mapping is used with 16 levels, the average absolute value of displacement due to a single

error is 3.75 levels. This figure fairly represents the average displacement per error over all errors (for p = 0.02 the probability of a single error is 0.075 and that of a double error is 0.0023).

Table VII summarizes the results of the simulation of data and application of the discrimination and correction technique. Points originally in error may either be improved (reduction or removal of error), worsened or left unchanged by the technique. In addition, some points not in error may suffer alteration. The sum total of the absolute value of deviation from the error-free value for all displaced data points after data-corrective processing has been completed is divided by the number of data points perturbed by simulated channel errors to obtain the average absolute value of displacement per channel error. The degree to which this value is less than the "standard" value of 3.75 is indicative of the benefit derived from the technique. The better performance obtained for A = 11 was with T = 4; a 40% reduction of average absolute displacement was obtained relative to straight binary mapping. Table VIII shows (A = 11, T = 4, 5) the distribution of displacements of those points which have been "corrected" by interpolation or extrapolation. Note that in both cases more than 70% of these points have displacements of two quanta or less.

The simulation demonstrates that modest gains in data reliability over a noisy channel may often be possible using the technique. Careful preliminary analysis of the data characteristics is a necessity to estimate the advantage which may be obtained.

## TABLE VII

### BREAKDOWN OF RESULTS OF DISCRIMINATION CORRECTION TECHNIQUE APPLIED TO SIMULATED DATA

| A | T | # of original errors in run | # of these errors | | | # of new errors added | av. Disp./orig. Error |
|---|---|---|---|---|---|---|---|
| | | | improved | worsened | unchanged | | |
| 11 | 4 | 39 | 33 | 0 | 6 | 3 | 2.28 |
| 11 | 5 | 39 | 29 | 0 | 10 | 2 | 2.51 |
| 15 | 5 | 37 | 28 | 0 | 9 | 3 | 3.11 |
| 15 | 6 | 37 | 24 | 0 | 13 | 2 | 3.52 |

TABLE VIII

TABULATION SHOWING NUMBER OF "DISPLACED" POINTS HAVING
A GIVEN DISPLACEMENT
(after interpolation)

~ Number of Points Having this Absolute Displacement

| Absolute Value of Displacement | A = 11 T = 4 | A = 11 T = 5 | Straight Binary Mapping |
|---|---|---|---|
| 0 | 6 | 5 | |
| 1 | 18 | 17 | ≈10 |
| 2 | 5 | 4 | ≈10 |
| 3 | 3 | 3 | |
| 4 | 3 | 4 | ≈10 |
| 5 | 3 | 2 | |
| 6 | 3 | 4 | |
| 7 | 1 | 2 | |
| 8 | | | ≈10 |

# REFERENCES

1. Heacock, R. L. et al., "Ranger VIII and IX, Part II, Experimenters Analyses and Interpretations," JPL Technical Report No. 32-800, 15 March 1966.

2. Widger, W. K., Jr., Meteorological Satellites, Holt, Rinehart and Winston, Inc., 1966.

3. Widger, W. K., Jr., et al., "Meteorological Interpretation of NIMBUS High Resolution Infrared (HRIR) Data," Allied Research Associates, Inc., NASA Contractor Report NASA CR-352, January 1968.

4. Anders, R. A. et al., "100 × 128-Element Solid State Imaging System," WESCON 67 Technical Paper 13/1, August 1967.

5. Lockheed Missiles and Space Co., "Asteroid Belt and Jupiter Flyby Mission Study," LMSC No. M-49-65-1 under JPL Contract No. 950871, February 1965.

6. Office of Naval Research Department of the Navy, "Handbook of Military Infrared Technology," edited by W. L. Wolfe, 1965.

7. Baghdady, E. J., "Topics in Telemetry Systems Analysis," ADCOM, Inc., Report No. 467-RR-76, 1967.

8. ADCOM, Inc., "Technical Report — Source Characterization Study," prepared under Contract NAS8-20001, 30 December 1966.

## Appendix A

## RUNLENGTH SIMULATION OF BINARY SYMMETRICAL CHANNEL

The purpose of this appendix is to describe the technique which was employed for the generation of simulated bit errors modeling the effects of data transmission over a binary symmetric channel. The brute force method of simulation would be to generate a uniformly distributed random number, $X$, in the range $0 \leq X < 1$ for each bit transmitted. If $X \leq p$ (where $p$ is the bit error probability) the bit in question is considered to be in error and its value is accordingly inverted. It is necessary to generate a random number for each bit.

It is clear that if $p \ll 1$, it would be far more efficient to generate a random variable which properly simulates the runlength between succesive errors than to apply error introduction on a bit-by-bit basis. The runlength generating procedure is outlined below.

Let $p(N)$ be used to denote the probability that the first error will occur at bit number $N$ with respect to the starting point of the data, or that an error will occur exactly $N$ bit intervals following the error immediately preceding. Thus, for a binary symmetric channel this function is:

$$p(N) = p \cdot (1 - p)^{N-1} \qquad (A-1)$$

Equation (A-1) above defines the probability of the runlength between errors assuming any integer value.

Visualize the division of the real line from 0 to 1 into an infinity of line segments. The length of the $k^{th}$ segment is simply $p(1 - p)^{k-1}$.
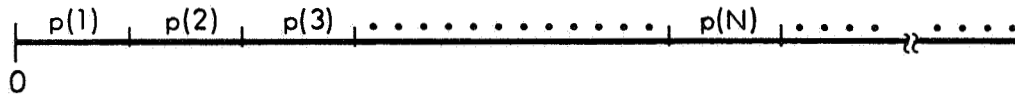


The runlength $N$ is generated if the uniformly distributed random number $X$ lies in the range:

$$\sum_{i=1}^{N-1} p(i) < X \leq \sum_{i=1}^{N} p(i) \qquad (A-2)$$

51

Equation (A-2) provides the means for mapping the random number X into the runlength N. To reduce this method to a form which may readily be implemented, it is necessary to get an expression for $S_N \equiv \sum\limits_{i=1}^{N} p(i)$:

$$S_N = p \cdot \sum_{i=1}^{N} (1 - p)^{i-1} = p \cdot \sum_{i=0}^{N-1} (1 - p)^N$$

$$= p \left[ \sum_{i=0}^{\infty} (1 - p)^i - \sum_{i=N}^{\infty} (1 - p)^i \right]$$

Noting that $\sum\limits_{i=0}^{\infty} (1 - p)^i = \dfrac{1}{p}$ :

$$S_N = 1 - (1 - p)^N \qquad\qquad (A-3)$$

In order for X to correspond to the value N, it is necessary that:

$$X \leq 1 - (1 - p)^N$$

which implies that:

$$\frac{\log(1 - X)}{\log(1 - p)} \leq N \qquad\qquad (A-4)$$

Also, since:

$$X > 1 - (1 - p)^{N-1}$$

we obtain:

$$N < \frac{\log(1 - X)}{\log(1 - p)} + 1 \qquad\qquad (A-5)$$

Substituting X for 1 - X (there is no difference on account of the uniform distribution) it is seen that a direct determination of N from X may be made by the relation:

$$N = \left[ \frac{\log X}{\log(1 - p)} \right] + 1 \qquad \text{(A-6)}$$

where [ ] denotes taking the integer part of the indicated quantity.

To implement the procedure, it is only necessary to generate a uniformly distributed random number over the range 0 to 1, take the log of this number to any convenient bases, divide by a precalculated constant, and add unity to the integer part of the quotient to produce the desired run-length.

Appendix B

# CONSTRUCTION AND ANALYSIS OF ANALOG-TO-DIGITAL MAPPING

Construction of single-error detecting A/D mapping. --We shall first describe the exact systematic procedure by which $2^\ell$ quantized signal levels are mapped into $2^\ell$ $\ell$-bit binary sequences in such a way that a single bit error in the binary sequence must correspond to at least a $2^{\ell-2}$ quantum level jump.

Let $C_{\ell-1}$ be used to denote the subset of all the $2^{\ell-1}$ $(\ell-1)$-tuples having an even number of 1's; the sequence of $\ell-1$ 0's is included in $C_{\ell-1}$. $C_{\ell-1}$ consists of exactly $2^{\ell-2}$ members. It is easily established that $C_{\ell-1}$ is a distance-2 code; i.e., at least two bits in any member of $C_{\ell-1}$ must be changed to transform it to any other member of $C_{\ell-1}$. $\overline{C}_{\ell-1}$ denotes the complement of $C_{\ell-1}$: those $(\ell-1)$-tuples of odd weight. $\overline{C}_{\ell-1}$ is also a distance-2 code. $C_{\ell-1}$, 0 and $\overline{C}_{\ell-1}$, 0 represent two sets of $\ell$-tuples containing $2^{\ell-2}$ members each formed by the appending of a 0 after each member of the set $C_{\ell-1}$ or $\overline{C}_{\ell-1}$. $C_{\ell-1}$, 1 and $\overline{C}_{\ell-1}$, 1 are similarly defined.

Figure B-1 shows how the mapping is constructed. The dotted line breaks up the signal range and corresponding sequence into quarters, of $2^{\ell-2}$ members each. The lower half of the listing of sequences is built up by using $C_{\ell-1}$, 0 and $\overline{C}_{\ell-1}$, 0. Any of the $(2^{\ell-2})$! possible orderings of members within each of these groups is permissible. Corresponding identical orderings of elements in $\overline{C}_{\ell-1}$ and $C_{\ell-1}$ are used for constructing the upper $2^{\ell-1}$ members of the listing.

To show that the desired error-displacement property is obtained by this mapping, we consider the effects of single errors upon a level in the second quarter of the scale, one whose $\ell$-bit representation is of the type $\overline{C}_{\ell-1}$, 1. No other member of $\overline{C}_{\ell-1}$, 1 is at a Hamming distance of unity from this sequence by virtue of $\overline{C}_{\ell-1}$ being a distance-2 code. No member of $C_{\ell-1}$, 0 is at unit distance from this sequence either: $C_{\ell-1}$ and $\overline{C}_{\ell-1}$ are disjoint sets, and the difference in the final digit ensures that no single error can transform a member of $\overline{C}_{\ell-1}$, 1 into a member of $C_{\ell-1}$, 0. The member of $\overline{C}_{\ell-1}$, 0 which is exactly $2^{\ell-2}$ levels above the sequence in question, must agree with it in the first $\ell-1$ places. The single error changing the final '1' of our sequence into a '0' will therefore lead to a shift upward of exactly $2^{\ell-2}$ levels. Note that no other member of $\overline{C}_{\ell-1}$, 0 can be produced with a single error.
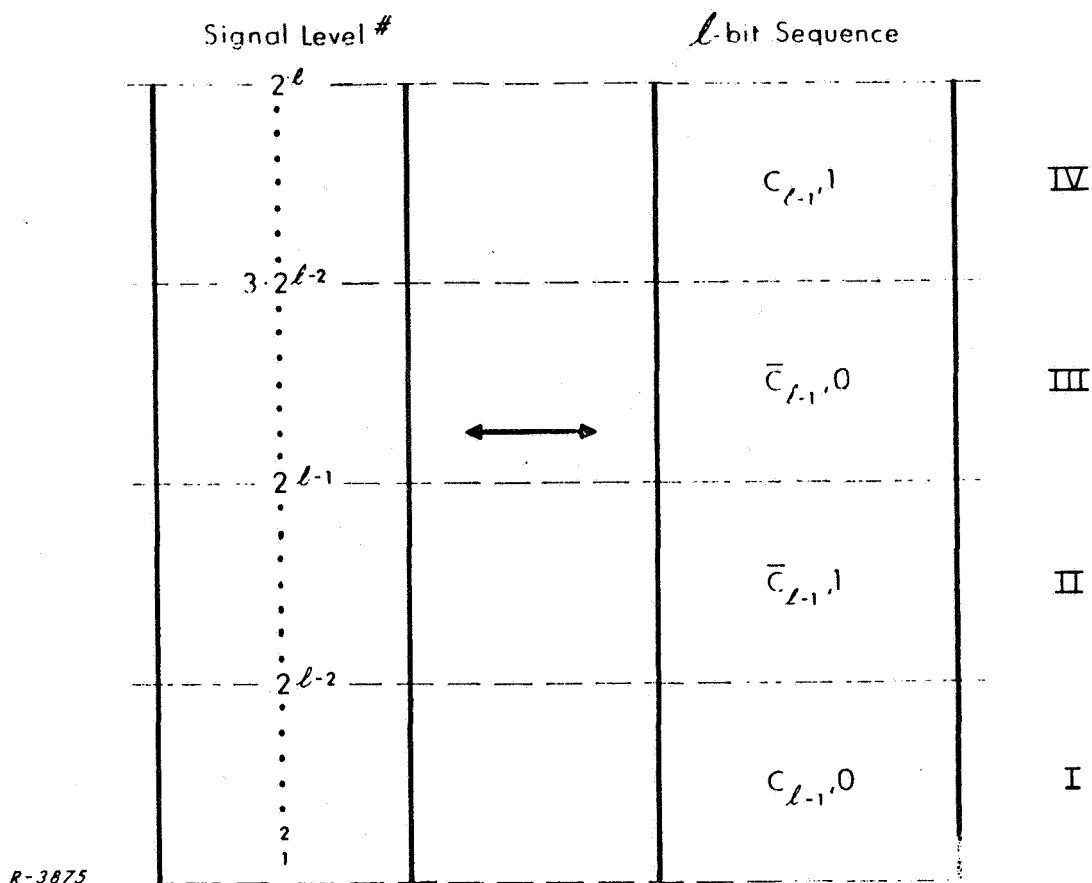
**Figure B-1   Construction of Analog-to-Digital Mapping**

Diagram labels: "Signal Level #" and "$\ell$-bit Sequence"

Signal levels from bottom to top: $2^1$, $2^{\ell-2}$, $2^{\ell-1}$, $3 \cdot 2^{\ell-2}$, $2^\ell$

$\ell$-bit sequences (bottom to top): $C_{\ell-1},0$ (region I), $\bar{C}_{\ell-1},1$ (region II), $\bar{C}_{\ell-1},0$ (region III), $C_{\ell-1},1$ (region IV)

R-3875

Average displacements for single and double errors. --An interesting characteristic of any analog-to-digital mapping technique is the manner in which bit errors in the digital sequence translate into displacements in gray scale level.   The evaluation of average displacements caused by single and by double errors has been carried out for two mappings: the single-error detecting mapping discussed above and straight binary mapping.

Single-error detecting mapping. --We had previously noted that this A/D mapping scheme guaranteed the mapping of single bit errors in the $\ell$-bit word into a level at least $2^{\ell-2}$ quanta away from the correct one. We now prove that the average displacement is $2^{\ell-1}$ quanta (i. e., half the total scale), and also establish a conservative approximation to the average displacement created by a double error.

56

Let us first consider single-error perturbations upon those levels which are members of grouping I in Figure B-1. An error in the $\ell^{th}$ bit will cause a shift up of exactly $3 \cdot 2^{\ell-2}$ quanta to the corresponding member in group IV. The other $\ell-1$ possible single errors must all transform a member of group I into a member of group III. Over all single errors applied to all members of group I, each member of group III will be counted an equal number of times. The average displacement created by single errors in the first $\ell-1$ places is simply the average level of group III minus the average level of group I.

Then the average displacement caused by a single error with respect to group I members is

$$\overline{D}_I = \frac{1}{\ell} \cdot 3 \cdot 2^{\ell-2} + \frac{\ell-1}{\ell} \cdot 2^{\ell-1} \qquad \text{(B-1)}$$

Similar reasoning leads to

$$\overline{D}_{II} = \frac{\ell-1}{\ell} \cdot 2^{\ell-1} + \frac{1}{\ell} \cdot 2^{\ell-2} \qquad \text{(B-2)}$$

$$\overline{D}_{III} = \frac{\ell-1}{\ell} \cdot 2^{\ell-1} + \frac{1}{\ell} \cdot 2^{\ell-2} \qquad \text{(B-3)}$$

$$\overline{D}_{IV} = \frac{\ell-1}{\ell} \cdot 2^{\ell-1} + \frac{1}{\ell} \cdot 3 \cdot 2^{\ell-2} \qquad \text{(B-4)}$$

Thus, the overall average is

$$\overline{D} = \frac{\overline{D}_I + \overline{D}_{II} + \overline{D}_{III} + \overline{D}_{IV}}{4} = 2^{\ell-1} \qquad \text{(B-5)}$$

In the evaluation of displacements due to double errors, one again begins by considering their effect upon group I members. A double error must either be one of the $\binom{\ell-1}{2}$ affecting only the first $\ell-1$ bits, or one of the $\ell-1$ in which the last bit is affected. The latter type, which represents the fraction $(\ell-1) \div \binom{\ell-1}{2} = 2/\ell$, transforms the affected sequence into a member of group II; the associated average displacement is $2^{\ell-2}$.

The other $\binom{\ell-1}{2}$ double errors must always generate sequences lying within group I. For lack of any recognized analytic technique for evaluating average displacements created by this class of error, it was decided to establish an upper bound to this average displacement. This was done by assuming that these $\binom{\ell-1}{2}$ double errors will act upon a member sequence of the group so as to produce those $\binom{\ell-1}{2}$ members of the group farthest away from the original sequence. These displacements must be averaged over all members of group I, and then given a weighting of $(\ell-2)/\ell$ to be combined with the average displacement produced by the other type of error. The results of this calculation are summarized in Table B-1. Note that both the results summarized in this table, and the single-error result of Eq. (B-5) above are completely independent of the many possible A/D mappings which may be produced by different permutations of the members of $C_{\ell-1}$ and $\overline{C}_{\ell-1}$.

## TABLE B-1
### AVERAGE DISPLACEMENTS DUE TO SINGLE AND DOUBLE ERRORS, OPTIMUM A/D MAPPING

| $\ell$ | Number of Levels | Average Displacement For | |
|---|---|---|---|
| | | Single Error | Double Error |
| 3 | 8 | 4 | $\leq 1.667$ |
| 4 | 16 | 8 | $\leq 2.833$ |
| 5 | 32 | 16 | $\leq 5.200$ |
| 6 | 64 | 32 | $\leq 10.333$ |

Straight binary A/D mapping. --Straight binary mapping refers to the scheme by which each level number is directly associated with its $\ell$-bit binary equivalent, e.g., for 32 levels the sequences 00000 through 11111 map into 0 to 31, respectively.

A single error in a sequence will cause a displacement of one level if it perturbs the least significant bit, two levels if it perturbs the next most significant bit, ... $2^{\ell-1}$ levels if it perturbs the most significant bit. The average displacement is simply

$$\overline{D} = \frac{1}{\ell}\left\{1 + 2 + \ldots + 2^{\ell-1}\right\} = \frac{2^{\ell} - 1}{\ell} \qquad (B-6)$$

A double error creates a numerical change whose absolute value must be of the form

$$\left| \pm 2^{q_1} \pm 2^{q_2} \right|$$

where $0 \leq q_2 < q_1 \leq \ell - 1$. If all transmitted levels are equiprobable, then 0 and 1 have equal probability of occurrence in any particular bit position. Therefore, half the time that a double error affects the two bit positions symbolizing $2^{q_1}$ and $2^{q_2}$, a displacement of absolute value $2^{q_1} + 2^{q_2}$ will be produced (two 1's changed to 0's, or two 0's changed to 1's), and the other half of the time the displacement will be of value $2^{q_1} - 2^{q_2}$. The average displacement associated with this particular error pattern affecting the digits symbolizing $2^{q_1}$ and $2^{q_2}$ is

$$\overline{D}(q_1, q_2) = \frac{2^{q_1} + 2^{q_2} + 2^{q_1} - 2^{q_2}}{2} = 2^{q_1} \qquad \text{(B-7)}$$

This displacement is independent of $q_2$. Note that there are, in general, $j$ double-error patterns having $q_1 = j$. The summation yielding $\overline{D}$ must be appropriately weighted to take account of this

$$\overline{D} = \frac{1}{\binom{\ell}{2}} \cdot \sum_{j=1}^{\ell-1} j \, 2^j \qquad \text{(B-8)}$$

Manipulation and subsequent evaluation of Eq. (B-8) yields

$$\overline{D} = \frac{4\{(\ell-2) \cdot 2^{\ell-1} + 1\}}{\ell(\ell-1)} \qquad \text{(B-9)}$$

Results are summarized in Table B-2.

## TABLE B-2
## AVERAGE DISPLACEMENTS DUE TO SINGLE AND DOUBLE ERRORS, STRAIGHT BINARY A/D MAPPING

| $\ell$ | Number of Levels | Average Displacement For | |
|---|---|---|---|
| | | Single Errors | Double Errors |
| 3 | 8 | 2.33 | 3.33 |
| 4 | 16 | 3.75 | 5.67 |
| 5 | 32 | 6.20 | 9.80 |
| 6 | 64 | 10.50 | 17.20 |
| 7 | 128 | 18.14 | 30.57 |

## Appendix C

## NEW TECHNOLOGY

After a diligent review of the work performed under this contract, no new innovation, discovery, improvement or invention was made.